International Journal of Current Research and Modern Education (IJCRME)
ISSN (Online): 2455 - 5428
(www.rdmodernresearch.com) Volume I, Issue I, 2016

THE ARTIFICIAL NEURAL NETWORK (ANN) BASED COMPUTATIONAL MODELLING FOR THE PREDICTION OF THE ACID-BASE PH-METRIC TITRATION VALUES

B. Thangagiri
Assistant Professor, Department of Chemistry, Mepco Schlenk Engineering College (Autonomous), Mepco Nagar, Sivakasi, Tamilnadu

Abstract:

In this study, a three-layered feed-forward artificial neural network (ANN) trained by back-propagation learning was used to model the complex non-linear relationship between the concentration of acid (HA) and the pH of solution at different volumes of the added titrant. The ANN developed model (2-13-1) with two input, one hidden and one output layers is used to predict the pH of the solution. The principal components of the pH matrix were used as the input of the network. The developed ANN model was trained with 70% of the data, validated with 15% data and tested with 15% remaining data set. The network architecture and parameters were optimized to give low prediction error. Using some of the experimental data for training, an ANN model based on standard Back-Propagation (BP) algorithm for the engine was developed. Multi layer perception network (MLP) was used for nonlinear mapping between the input and the output parameters. The different activation functions and several rules were used to assess the percentage error between the desired and the predicted data set. The optimised networks predicted the pH of the different concentrations of acid-base mixtures. It was observed that the ANN model can predict the pH quite well with correlation coefficient (R) was 0.99987. The prediction MSE (Mean Square Error) error was between the desired outputs as measured values and the simulated values by the model were obtained as 0.0001.

Key Words: pH-Metric Titration, Acid-Base Titration & Artificial Neural Network.

1. Introduction:

The pH control is very important in many processes such as in wastewater treatment plant, the cell growth rate and the accurate stabilisation of pH at an optimal level often determines the efficiency of the bioprocess. The regulation and control of a pH process is a typical problem found in a variety of industries including wastewater treatment, pharmaceuticals, biotechnology and chemical processing. It is a nontrivial task arising from the nonlinearity of the titration process. Therefore, controlling the pH at certain region or set point is very important. On the other hand, in chemical processes, pH neutralization is not easy to control due to the fast and quite complicated reaction [1-2]. An acid–base titration is the determination of the concentration of an acid or base by exactly neutralizing the acid or base with an acid or base of known concentration. This allows for quantitative analysis of the concentration of an unknown acid or base solution. A good model would reduce the necessity for elaborate measurement procedures which is expensive and time consuming. In this case, deterministic models are difficult to develop owing to the complexity of emissions formation mechanism. Therefore, a non-linear modelling approach (ANN) was adopted for modelling and prediction. In recent years, ANNs are increasingly being used to solve engineering problems that deal with highly nonlinear functional approximations [3].

Artificial neural networks (ANNs) are biologically inspired computer programs designed to simulate the way in which the human brain processes information. ANNs gather their knowledge by detecting the patterns and relationships in data and learn (or are trained) through experience, not from programming. Artificial neural networks (ANNs) are now commonly used in many areas of chemistry and they represent a set of methods that may be useful in solving many problems [4-5]. ANNs are directly inspired from the biology of the human brain, where billions of neurons are interconnected to process a variety of complex information. Accordingly, a computational neural network consists of simple processing units called neuron. Each network consists of artificial neurons grouped into layers and put in relation to each other by parallel connections. The strength of these interconnections is determined by the weight associated with them. For every ANN, the first layer constitutes the input layer (independent variables) and the last one forms the output layer (dependent variables). Between them, one or more neurons layers called hidden layers can be located. The number of input and output neurons effectively represents the number of variables used in the prediction and the number of variables to be predicted, respectively. The hidden layers act as feature detectors and, in theory, there can be more than one hidden layer. The universal approximation theory, however, suggests that a network with a single hidden layer. With a sufficiently large number of neurons can interpret any input and output structure [6].

ANNs are used to solve a wide variety of problems in science and engineering, particularly for some areas where the conventional modelling methods [7]. The various applications of ANNs can be summarised into classification or pattern recognition, prediction and modelling. Supervised associating networks can be applied in pharmaceutical fields as an alternative to conventional response surface methodology. Unsupervised feature-extracting networks represent an alternative to principal component analysis. Non-adaptive unsupervised networks are able to reconstruct their patterns when presented with noisy samples and can be used for image recognition.

A well trained ANN can be used as a predictive model for a specific application, which is a data-processing system inspired by biological neural system. The predictive ability of an ANN results from the training on experimental data and then validation by independent data. An ANN has the ability to re-learn to improve its performance if new available data [8]. An ANN model can accommodate multiple input variables to predict multiple output variables. It differs from conventional modelling approaches in its ability to learn about the system that can be modelled without prior knowledge of the process relationships. The prediction by a well-trained ANN is normally much faster than the conventional simulation programs or mathematical models as no lengthy iterative calculations are needed to solve differential equations using numerical methods but the selection of an appropriate neural network topology is important in terms of model accuracy and model simplicity. In addition, it is possible to add or remove input and output variables in the ANN if it is needed. The objective of this study was to develop a neural network model for predicting pH values in relation to input variables such as volume of hydrochloric acid and sodium hydroxide. This model is of a great importance due to its ability to predict performance under varying conditions

2. Artificial Neural Network (ANN) Modelling:

ANN is an approach inspired by brain structure and tries to simulate the brain processing capabilities. Neural network operates like a "black box" model, and does not require detailed information about the system. Instead, it learns the relationship between the input parameters and the controlled and uncontrolled variables by studying previously recorded data, in a similar way that a non-linear regression might be performed. A neural network is composed of large numbers of highly interconnected processing elements known as neurons. Basically, a biological neuron receives inputs from other sources, combines them in some way, performs generally a non-linear operation on the result, and then outputs the final result. The network usually consists of an input layer, some hidden layers, and an output layer. Information from the input layer is processed in the hidden layer. The hidden layer neurons then process the incoming information and extract useful features to reconstruct the mapping from the input space. The neighbouring layers are fully interconnected by weights. Finally, the output layer neurons produce the network prediction to the outside world.

3. Material and Methods:

3.1 Materials:

All the reagents used were of analytical grade. Double distilled water was used throughout. Stock solutions were prepared by dissolving the appropriate amounts of NaOH and HCl in water.

3.2. General Procedure:

The pH-metric titrations were carried out in a double walled cell, and the temperature was kept constant using a water bath circulating system at a constant temperature of 25 °C. The pH of the solution was measured by a digital pH-meter (Systronics, model 335) using a combined glass electrode as shown in figure–1.

Figure 1: The pH-meter for acid-base titration

To the 20 mL of 0.01N hydrochloric acid with 20 mL of double distilled water, 0.4 mL of 0.0036N sodium hydroxide was added sequentially. For each addition of 0.4 mL of sodium hydroxide with proper stirring, the corresponding pH values were recorded using combined glass electrode. The results with ANN predicted values are shown in table–1 and the pH-metric graph is shown in figure–2.

Table 1: Experimental and ANN predicted pH values

Volume of NaOH added (mL)	Experimental pH	ANN predicted pH
0	1.91	1.91
0.4	1.94	1.94
0.8	1.97	1.97
1.2	2.01	2.01
1.6	2.05	2.05
2	2.11	2.11
2.4	2.16	2.16
2.8	2.22	2.22

3.2	2.29	2.29
3.6	2.37	2.37
4	2.47	2.47
4.4	2.59	2.59
4.8	2.76	2.76
5.2	3.05	3.05
5.6	7.47	7.47
6	10.3	10.3
6.4	11.22	11.22
6.8	11.59	11.59
7.2	11.75	11.75
7.6	11.85	11.85
8	11.93	11.93
8.4	11.99	11.99
8.8	12.05	12.05
9.2	12.09	12.09
9.6	12.13	12.13
10	12.16	12.16

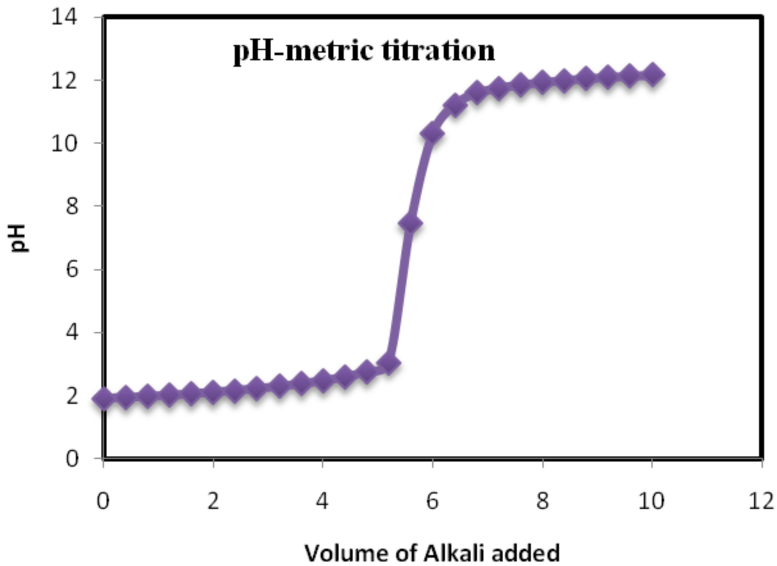

Figure 2: The pH-metric graph

3.3. ANN Software:

All ANN calculations carried out using Matlab 8.1 (R2013a) mathematical software with ANN toolbox for windows running on a personal computer (Pentium IV 2800 MHz). A three-layer network with a sigmoidal transfer function (trainscg) with back propagation algorithm was designed in this study [9-10].

3.4. Design of Artificial Neural Network:

Artificial neural networks (ANN) are computer systems developed to have such abilities as automatically generating, forming and discovering new knowledge without any help just as the human brain [11]. Generally, it is composed of three types of layers; an input layer, a few hidden layers and an output layer [12]. Each layer has a certain number of components called neurons or nodes which are linked to each other. Each neuron is linked to others with communication links accompanied by linking weight.

The signals pass through neurons on the linking weights. Each neuron takes multiple inputs from other neurons in accordance with their linking weights and may generate an output signal which can also be generated by other neurons [13]. The structure of ANN is shown figure–3. To develop an ANN model, the network was subjected to two processes; training and test. In training, the network is trained to estimate output values relative to input data. In testing, the network is tested to stop training or save training data and is used to estimate an output.

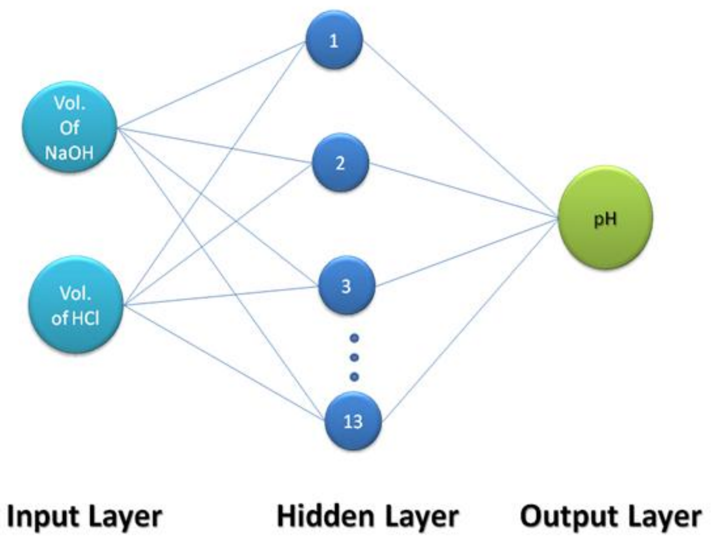

Figure 3: The selected neural network structure

When the tested error reaches previously determined tolerance value, the training process is finished [14]. The topology of an artificial neural network (ANN) is determined by the number of layers, the number of nodes in each layer and the nature of the transfer functions. Optimization of ANN topology is probably the most important step in the development of a model [15].

Back propagation (BP) algorithm is the most popular and most commonly used algorithm. BP is composed of two phases; feed forwarding and back propagation. Knowledge subjected to processing from the input layer up to the output layer is generated during feed forwarding. In back propagation knowledge subjected to processing from the input layer up to the output layer is generated during feed forwarding. In back propagation phase, the difference between network output value obtained in feed forwarding and desired output value is compared with previously determined difference tolerance and the error in output layer is calculated. This error value is propagated backward to update the links in the input layer [15]. The BP training algorithm is a ramp descent algorithm. BP algorithm minimizes total error by changing the weights through its ramp and thus tries to improve the performance of the network. The training of the network is stopped when the tested values of mean square error (MSE) stop decreasing and begin to increase. The estimated performance is calculated by using the equation–1.

$$MSE\% = \frac{1}{n}\sum_{i=1}^{n}(d_i - O_i)^2 \qquad (1)$$

where d is the aimed or real value, O is network output or estimated value, n is the number of the output data. In this work, the sigmoidal transfer function was used as a transfer function in the hidden and output layers. This is the most widely used transfer function, which is given by using the equation–2 where $f(x)$ is the hidden neuron output.

$$f(x) = \frac{1}{1+\exp(-x)} \quad (2)$$

3.5. Modelling with Artificial Neural Network:

The aim in designing this ANN is to estimate the performance depending on the amount of input and output particles in relation. The data set obtained from experimental study was divided into two as training and test data sets.

In total, 26 experimental sets were used to feed the ANN structure. The data sets were divided into training, validation and test subsets, each of which contains 18 (70% of total data), 4 (15% of the total data) and 4 (15% of the total data) samples, respectively. The validation and test sets, for the evaluation of the validation and modelling power of the nets, were randomly selected from the experimental data.

In this study, a network with feed forward and with one input layer, one hidden layer and four output layer was used as shown in figure–5. All samples were normalized in the 0–1 range. So, all of the data sets (Xi) (from the training, validation and test sets) were scaled to a new value xi as follows by using the equation–3.

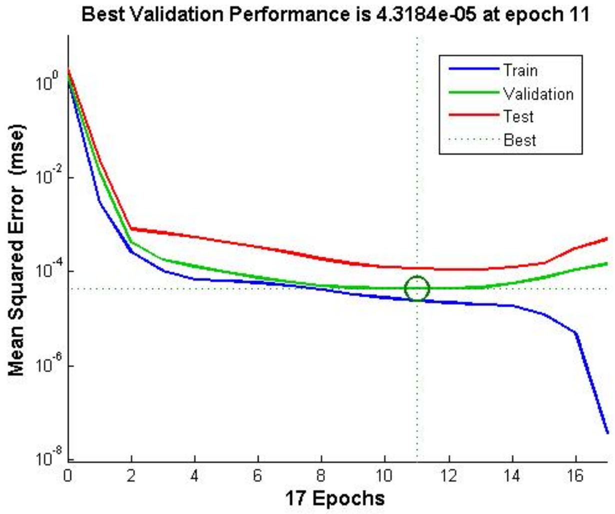

Figure 4: The best validation performance of the network at epoch=11

$$X_i = \frac{X_i - X_{min}}{X_{max} - X_{min}} \quad (3)$$

After the ANN structure was developed, the data set obtained in experimental study was normalised in a set containing all values between 0 and 1 to improve training characteristic. In training process, the back-propagation algorithm was used. As the hidden and output layers are composed of non-linear processing units as neurons, Tangent–Sigmoid (TANSIG) transfer function[16] was chosen in this study. The table–2 shows the performance of ANN model for the prediction of pH values. Based on low MSE value the optimum ANN model with 13 neurons in hidden layer is selected. Hence, the optimum performance is shown by 2-13-1 ANN model which is shown in figure-3.

Table 2: The performance of ANN model for the prediction of pH values

No of Neurons	MRE	R
3	0.0000944485	0.99703
4	0.000809956	0.99754
5	0.00023533	0.9993
6	0.000345558	0.99887

7	0.0001571999	0.99947
8	0.0000754502	0.99975
9	0.000469777	0.99986
10	0.000129334	0.99958
11	0.00154479	0.99491
12	0.000118477	0.99962
13	0.0000409105	0.99987
14	0.000148842	0.99162
15	0.000365928	0.99886
16	0.0010244	0.99664
17	0.000279488	0.99913
18	0.00352479	0.98825
19	0.00891869	0.97645
20	0.0414498	0.91419

Comparisons of the ANN predictions and experimental results for testing sets of output performance parameters are given in table–1 for easy comparisons of input parameters. The figure–4 shows that the best validation performance obtained at epoch is 11. The regressions analysis of training, validation and testing are shown in figure–5.

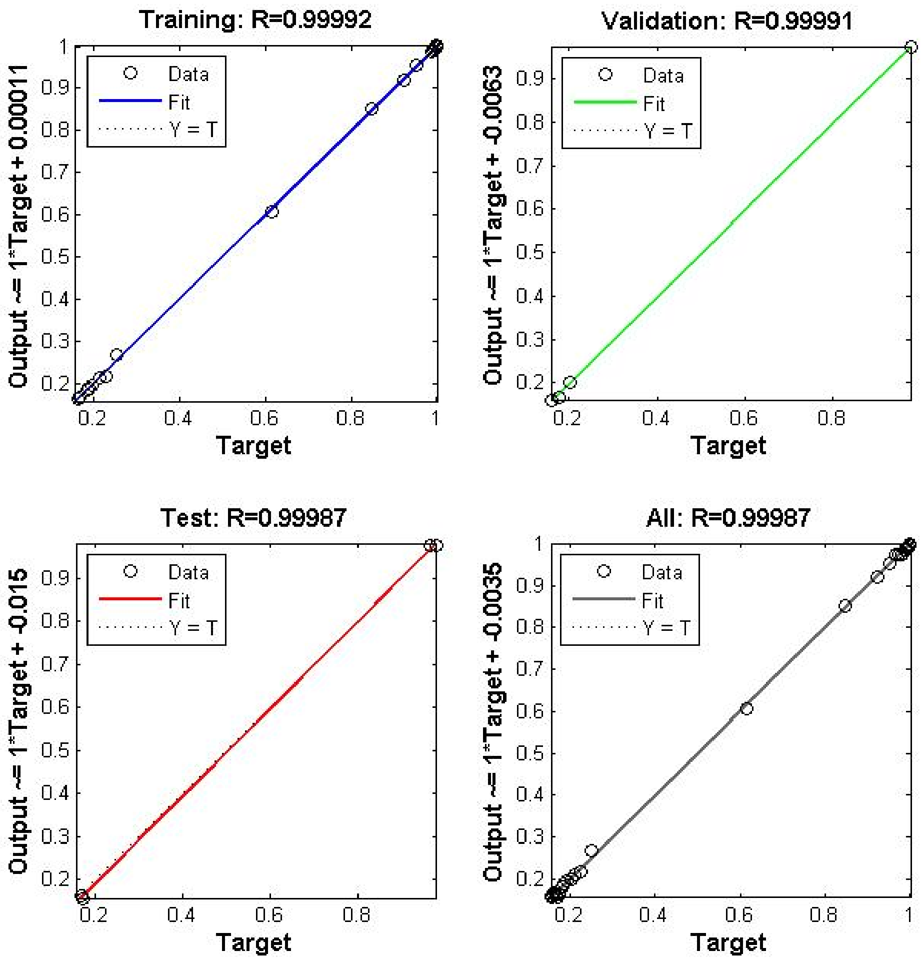

Figure 5: Regression analysis of Training, Validating and Testing of network

4. Results and Discussion:

The use of an ANN model is considered as a practical and reliable approach for non-linear problems like pH-metry experiments. The ANN predicted outputs are de-normalised and given in the same table–1.

The test pattern of experimental pH values and ANN predicted pH values are given in figure–6. Thefigure–7 shows that there is good correlation analysis between experimental and ANN predicted pH values. It shows that the ANN modelling can predict the pH values very accurately with correlation co-efficient of 0.9999 and Mean Square Error (MSE) value of 0.0001 and RMSE value is also very low. Hence, ANN modelling of pH-metric titration can give the following information.

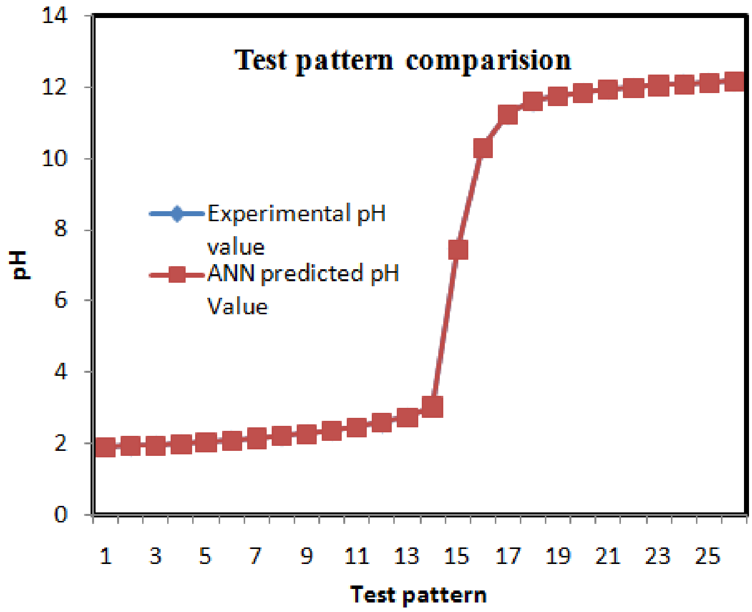

Figure 6: Test pattern of experimental and ANN predicted pH values

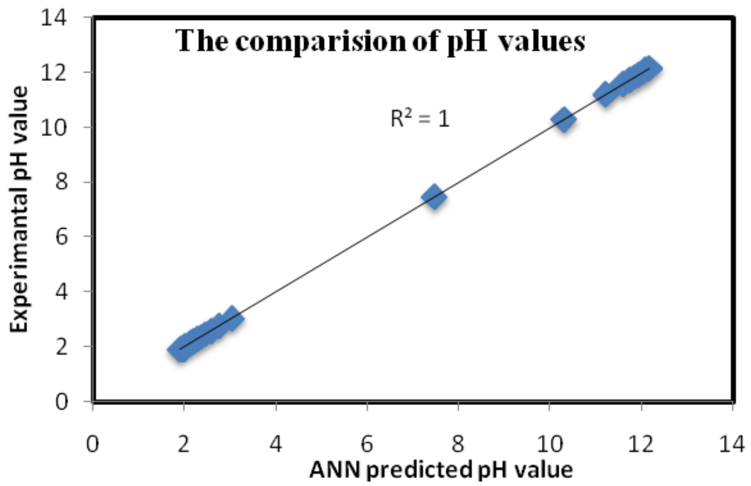

Figure 7: The correlation between the Test pattern of experimental and ANN predicted pH values

1. No need of background of the input data. It can be used for non-liner curve fitting like pH-metric titration.
2. These saves time, chemicals and make eco-friendly Green Chemistry through green computational techniques.

3. It allows us to face the above problem without much knowledge of Chemistry concept like pH.
4. This may have advantage only if a problem cannot be solved by clear algorithm.

5. Conclusion:

The applicability of ANN model (2-13-1) has been investigated for the performance of prediction of pH values of different concentrations. To train the network, the volume of hydrochloric acid and sodium hydroxide given as input layer while ANN predict the pH, as the required output. By using the back-propagation learning algorithm with three different variants, single layer, and logistic sigmoid transfer function, the weights of the network, formulations have been given for each output. The network has yielded correlation co-efficient (R^2) values of 0.99992 and the mean % errors are smaller than 0.001 for the training data, while the R^2 values are about 0.99987 and the mean % errors are smaller than 0.01 for the test data. The results may easily be considered to be within the acceptable limits. The relationship between volume of hydrochloric acid and sodium hydroxide with pH can be determined for different concentration of solution by using the network. Therefore, the usage of ANNs may be highly recommended to predict the pH values of the solution instead of having to undertake instruments and time-consuming experimental studies. Simulation based on the ANN model can then be performed in order to estimate the behaviour of the computational neural network system under different conditions. Analysis of the experimental data by the ANN revealed that there was good correlation between the predicted data resulted from the ANN and measured ones. The developed model thus reduces the experimental efforts and hence can serve as an effective tool for predicting the pH under various concentrations. This information is essential for the adequate scale-up and design of industrial scale batch reactors for the treatment of organic contaminants in wastewaters as well as waste water treatment process in developing countries.

6. Acknowledgement:

Researchers express their gratitude to the Principal and Management of Mepco Schlenk Engineering College (Autonomous), Sivakasi for their constant support and encouragement.

7. References:

1. H.C. Chan and C.C. Yu, Autotuning of gain scheduled pH control: An experimental study, Industrial and Engineering Chemistry Research, 34 (1995) 1718-1729.
2. D.J. Costello, Evaluation of model based control techniques for a buffered acid-base reaction system, Process operation and control, Chemical Engineering Research and Design. 72 (1994), 47-54.
3. C.A.O. Nascimento, R. Giudici and R. Guardani, Neural network based approach for optimization of industrial chemical processes, Computers and Chemical Engineering. 24 (2000) 2303–2314.
4. J.A. Stegemann and N.R. Buenfeld, Prediction of leachate pH for cement paste containing pure metal compounds, Journal of Hazardous Materials, B90 (2009) 169-188.
5. V.K. Pareek, M.P. Brungs, A.A. Adesina and R. Sharma, Artificial neural network modelling of a multiphase photo-degradation system, Journal of Photochemistry and Photobiology A: Chemistry. 149 (2000) 139-146.
6. F. Despange and D.L. Massart, Neural networks in multivariate calibration, Analyst. 123 (1998) 157-178..

7. Johann Gasteiger and Jure Zupan, Neural networks in Chemistry, Angewandte Chemie International Edition in English. 32 (1993) 503-521.
8. B. Ghobadian, H. Rahimi, A.M. Nikbakht, G. Najafi and T.F. Yusaf, Diesel engine performance and exhaust emission analysis using waste cooking biodiesel fuel with an artificial neural network, Renewable Energy. 34 (2009) 976-982.
9. Hamid Taghavifar and Aref Mardani, Application of artificial neural networks for the prediction of traction performance parameters, Journal of the Saudi Society of Agricultural Sciences. 13 (2014) 35-43.
10. L. Tang, G.M. Zeng, G.L. Shen, Y. Zhang, G.H. Huang and J. B.Li, Simultaneous amperometric determination of lignin peroxidase and manganese peroxidase activities in compost bioremediation using artificial neural networks, Analytica Chimica Acta. 579 (2006) 109-116.
11. M. Canakci, A.N. Ozsezen, E. Arcaklioglu and A. Erdil, Prediction of performance and exhaust emissions of a diesel engine fueled with biodiesel produced from waste frying palm oil, Expert systems with Applications. 36 (2009) 9268-9280.
12. G.G. Lenzi, R.F. Evangelista, E.R. Duarte, L.M.S. Colpini, A.C. Fornari, R. Menechini Neto, L.M.M. Jorge and O.A.A. Santos, Photocatalytic degradation of textile reactive dye using artificial neural network modeling approach. Desalination and Water Treatment, DOI: 10.1080/19443994.2015.1064035 (2015).
13. C.G. Looner, Pattern Recognition using Neural Networks, Theory and Algorithms for Engineers and Scientists, Oxford University Press, New York, pp. 75-78 (1997).
14. S.A. Kalogirou, Artificial neural networks in the renewable energy systems applications: a review, Renewable and Sustainable Energy Reviews. 5 (2001) 373–401.
15. Guangying Chen, Xiao Luo, Haiyan Zhang, Kaiyun Fu, Zhiwu Liang, Wichitpan Rongwong, Paitoon Tontiwachwuthikul and Raphael Idem, Artificial neural network models for the prediction of CO_2 solubility in aqueous amine solutions, International Journal of Greenhouse Gas Control. 39 (2015) 174-184.
16. Hossein Kamran Haghighi, Meraj Rafie, Davood Moradkhani, Behzad Sedaghat and Aliakbar Abdollahzadeh, Modeling on Transition of Heavy Metals from Ni–Cd Zinc Plant Residue Using Artificial Neural Network, Transactions of the Indian Institute of Metals. DOI: 10.1007/s12666-014-0507-3 (2015).